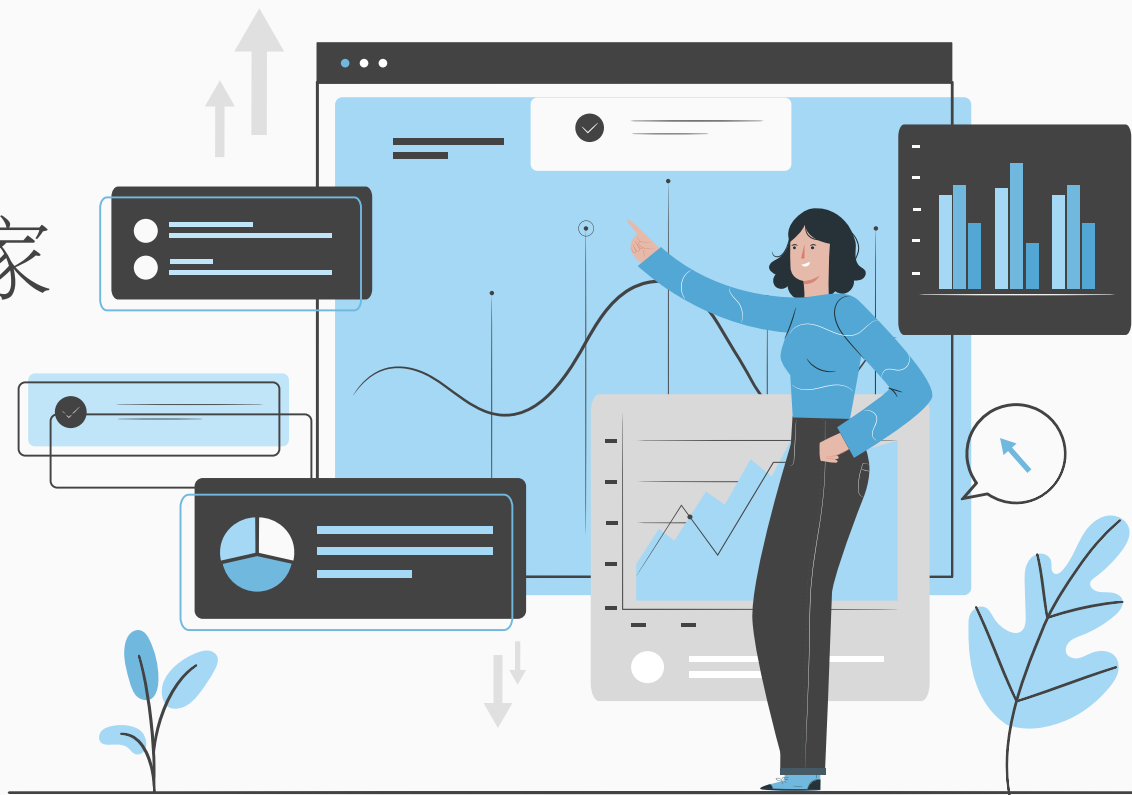


# 個人投資風險管家 玉山證券

台大財經 林鈺晉  
東吳巨資 周采薇  
李悅暄  
葉芯好



專案廣告



01

專案背景

02

抽樣條件  
特徵挑選

03

**RANDOM FOREST**  
隨機森林

04

**SVM** 支援向量機

05

模型比較  
遇到困難

06

問卷成果  
**CHATBOT DEMO**

# 01. 專案背景





## 專案背景



觀察違約交割的案例後，發現部分案例發生在投資新手，初期投資方式較為保守，但在幾次的獲利後，對風險偏好上升卻忽略了自身承受風險的能力，因此讓後續投資失利並且無法如期交割。

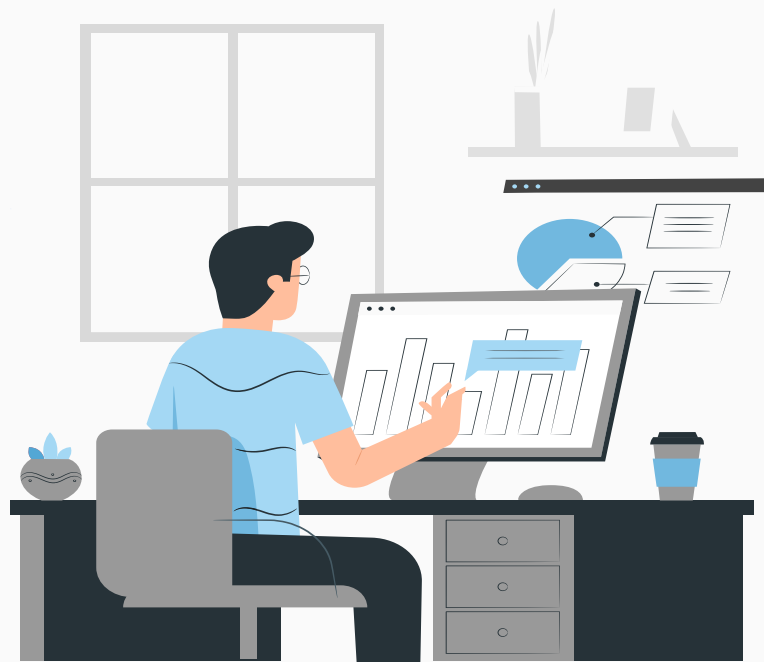
因此希望能夠偵測出投資人的投資行為改變，找出投資人的異常行為，進一步判斷是否反映出風險偏好的改變，並透過簡訊的方式通知投資人。

**目前市面上的做法：**開戶或提出提高交易額度申請時的審查制度，再透過交易額度上限避免投資人過度投資。

**本組作法：**因為部分的違約案例發生在投資新手，因此我們把目標定為投資交易經驗2年以下。而違約交割屬於少數異常的行為，因此以找出異常行為為主。



## 02. 抽樣條件 特徵挑選



## 抽樣條件

01

違約人年齡區間

AGE_LEVEL	
2	16
3	14
4	10
6	3
5	1
1	1

違約人的年齡區間中以2、3、4為最多，因此抽樣時以年齡區間中以2、3、4為主。

02

新手投資人

0: 未滿一年  
1: 1~2 年  
2: 2~3 年  
3: 3~5 年  
4: 5~10 年  
5: 10 年以上

本組決定針對新手投資人進行分析，因此抽樣時以投資經驗0、1為主

03

時間[季度]

2019 Q1 2020 Q1 2021 Q1  
2019 Q2 2020 Q2  
2019 Q3 2020 Q3  
2019 Q4 2020 Q4

本題因包含時間序列，需要用過去資訊預測未來，因此抽樣時以每一季度抽取2500筆為主

### 市場面

- **PL**：當天損益  
[(收盤價格-交易價格)/交易價格]
- **Amplitude**：交易當天最大振幅  
[(最高價-最低價)/開盤價]
- **INDUSTRY**：產業別
- **CAPITAL\_TYPE**：公司規模(大中小)
- **VOLUME**：交易股數(每日股票資訊)
- **VOLUME**：交易股數(每日股票資訊)
- **AMONT**：交易量
- **ROI**：報酬率
- **ALPHA**：ALPHA值
- **BETA\_21D**：BETA值(月)
- **BETA\_65D**：BETA值(季)

### 交易面

- average\_30：過去30日平均交易量
- average\_10：過去10日平均交易量
- STOCKS：交易股數(交易紀錄)
- COMMISSION\_TYPE\_CODE：交易別
- MARKET\_TYPE\_CODE：市場別

### 客戶面

- BRANCH\_IND：違約註記
- BUY\_COUNT：交易次數(買入)
- SELL\_COUNT：交易次數(賣出)
- NONTXN\_COUNT：交易次數(非交易)



### 03.

# RANDOM FOREST

## 隨機森林



違約人年齡區間2~3

## RANDOM FOREST

### 01 MODEL

```
clf = RandomForestClassifier(n_estimators = 40, max_depth = 8)
```

### 02 SCORE

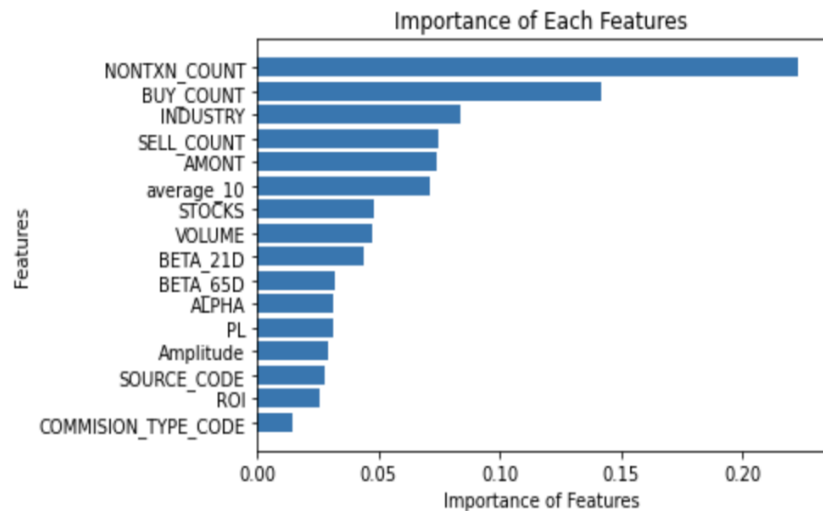
	Precision	Recall	F1_Score
Train	99.92%	82.78%	89.31%
Validation	99.89%	75%	83.04%
Test	92.06%	52.93%	55.43%

### 03 FEATURES SELECT

Delete features:

CAPITAL\_TYPE：公司規模(大中小)

MARKET\_TYPE\_CODE：市場別



違約人年齡區間3~4

## RANDOM FOREST

### 01 MODEL

```
clf = RandomForestClassifier(n_estimators = 40, max_depth = 8)
```

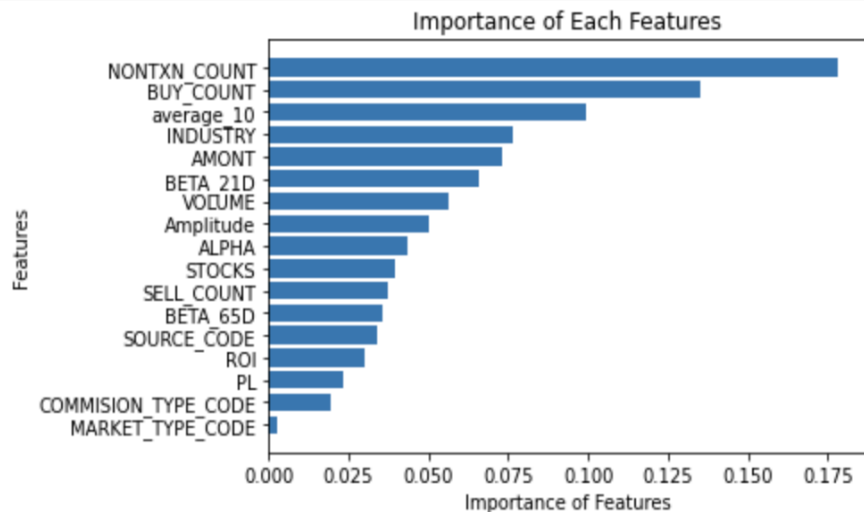
### 02 SCORE

	Precision	Recall	F1_Score
Train	99.92%	82.47%	89.32%
Validation	99.89%	74.7%	82.99%
Test	99.58%	53.94%	57.07%

### 03 FEATURES SELECT

Delete features:

CAPITAL\_TYPE：公司規模(大中小)



# 04. SVM 支撐向量機



違約人年齡區間2~3

## 01 MODEL

```
clf = SVC(C=250 , kernel='rbf', gamma='auto')
```

## 02 SCORE

	Train	Validation	Test
F1_Score	69.72%	66.45%	57.29%

## 03 FEATURES SELECT

Delete features by Backward Stepwise:

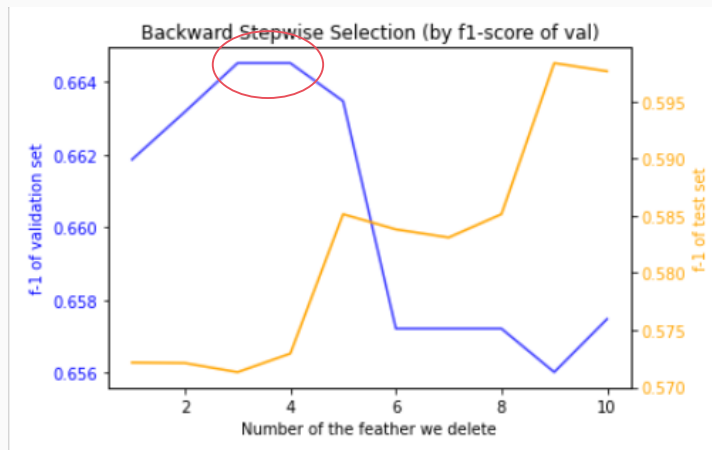
Sell\_COUNT : 交易次數(賣出)

BUY\_COUNT : 交易次數(買入)

ALPHA : ALPHA值

PL : 當天損益

# SUPPORT VECTOR MACHINE



違約人年齡區間3~4

## 01 MODEL

```
clf = SVC(C=250 , kernel='rbf', gamma='auto')
```

## 02 SCORE

	Train	Validation	Test
F1_Score	69.54%	65.60%	61.16%

## 03 FEATURES SELECT

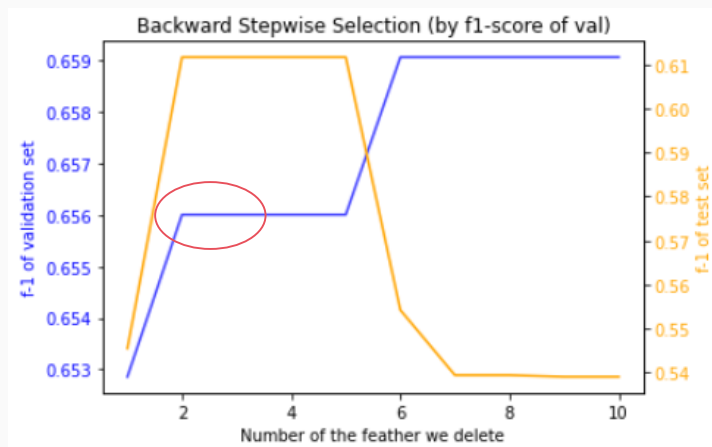
Delete features by Backward Stepwise:

Amplitude : 當天最大振幅

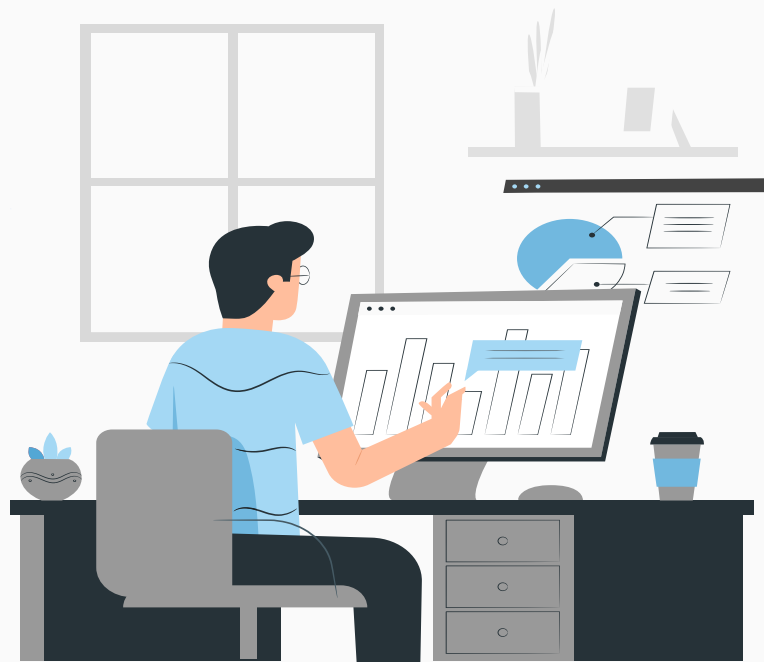
Amount : 交易量

Capital Type : 公司規模(大中小)

# SUPPORT VECTOR MACHINE



## 05. 模型比較 遇到困難



## 模型比較

### OI MODEL COMPARED

違約人年齡 區間2~3	Train	Validation	Test
RF	89.31%	83.04%	55.43%
SVM	69.72%	66.45%	57.29%

違約人年齡 區間3~4	Train	Validation	Test
RF	89.32%	82.99%	57.07%
SVM	69.54%	65.60%	61.16%

Random Forest在 Validation Set 有較好的f1-score，但Test Set卻比SVM低。這裡推測是因為random forest本身模型特性就比較容易overfitting.

乍看之下，這兩個model的testing f1-score 的分數都沒有很高，但仔細去查看predict value後發現，其實錯誤的地方幾乎都是在違約交割的data上判斷錯誤，而非違約交割的data上卻幾乎沒有判斷錯誤。

BUT WHY?



## 02 DIFFICULTY

### 1. 違約交割的比例極低

- 約僅有100左右的違約資料
- 違約資料沒有明顯的特殊性
- Random Forest overfitting 機率高
- SVM容易忽略違約資料

### 2. 特徵篩選困難

- 加入金融上可以拿來判斷風險的指標當作特徵
- 從Random Forest的feature importance看起來這些指標似乎不太重要。
- 在訓練SVM時做Backward Stepwise Selection時，發現刪除某些金融指標反而能夠提升準確度。

## 06. 問卷成果 CHATBOT

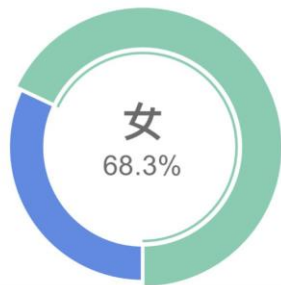


# 問卷成果

填答人數：231 / 231人

## 1 性別

答案選項	填答次數	百分比
女	158	68.3%
男	73	31.6%



填答人數：231 / 231人

## 2 年齡

答案選項	填答次數	百分比
20-30	84	36.3%
51-60	68	29.4%
41-50	39	16.8%
31-40	28	12.1%
60以上	12	5.1%



填答人數：231 / 231人

## 3 是否曾經有過投資行為

答案選項	填答次數	百分比
是	138	59.7%
否	93	40.2%



## 問卷成果

以下問題，請以假設您的投資行為可能產生較高的風險狀況回答

填答人數：231 / 231人

☒ 5 您希望證券公司以什麼方式提醒您？

答案選項	填答次數	百分比
<input checked="" type="checkbox"/> 簡訊	122	52.8%
<input type="checkbox"/> Email	109	47.1%
<input type="checkbox"/> 聊天軟體(例如：Line等)	88	38%
<input type="checkbox"/> 電話	60	25.9%

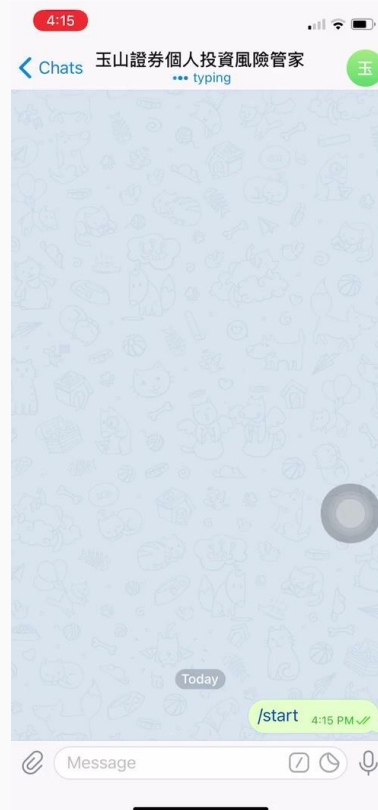
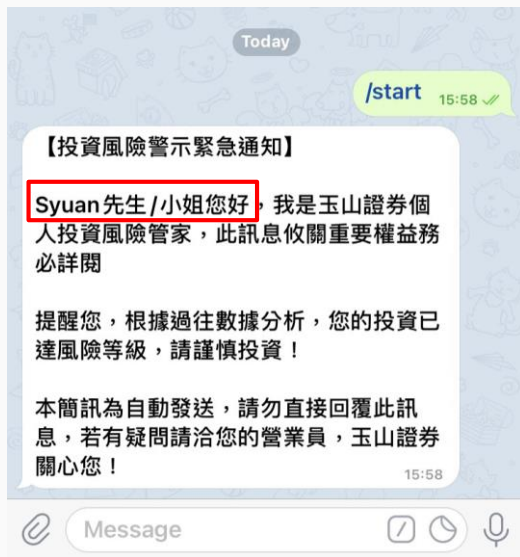
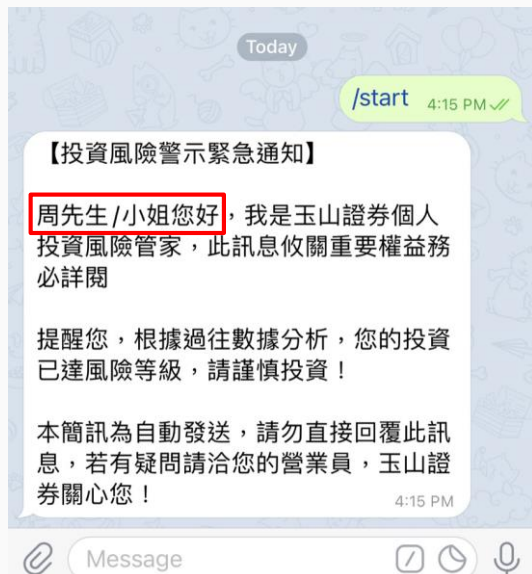
填答人數：231 / 231人

☒ 6 您希望證券公司於什麼時間提醒您？

答案選項	填答次數	百分比
<input checked="" type="checkbox"/> 當天收盤後(13:30-17:00)	152	65.8%
<input type="checkbox"/> 當天傍晚(17:00~21:00)	72	31.1%
<input type="checkbox"/> 隔天上午(08:00~12:00)	34	14.7%
<input type="checkbox"/> 當天深夜(21:00之後)	17	7.3%
<input type="checkbox"/> 隔天下午(12:00~17:30)	7	3%

## CHATBOT DEMO

因此，本組決定以Telegram Chatbot的形式模擬簡訊發送風險警示訊息



## 警示訊息文案建議：

- 提醒此為重要訊息
- 簡單易懂
  - 即時
  - 重點資訊、股票關鍵字
- 提供詳細資訊作為被判定為高風險的依據
  - 投資人的投資標的相關資訊及歷史資訊
  - 投資人投資行為歷史數據分析
  - 相關的大盤資訊

# THANK YOU FOR YOUR PATIENCE

**CREDITS:** This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik** and illustrations by **Storyset**



專案廣告

